Political Persuasion in the Information Age: Results from a Field Experiment

Raymond Fisman* Yiying Li† Meng Miao‡

August 11, 2025

Abstract

We study the effectiveness of online persuasion via a lab-in-the-field experiment that exposes respondents to short videos on the proposed Public Security Law in China. Exposure to "nationalist" videos supporting the law reduced willingness to provide oppositional feedback to the government by 4.8 percentage points (a 42% reduction relative to the control), while critical videos increased oppositional feedback by 6.6 percentage points (58% relative increase); for both nationalist and critical videos, the effect on providing supportive feedback was the opposite. In a follow-up experiment, we informed subjects of the extent to which others had a negative view of those criticizing the law, which completely erased any effect of the video on willingness to provide feedback. Overall, our findings indicate the potency of online content in its short-run impact on political dissent, and that these effects may operate in large part through their impact on second-order beliefs.

Keywords: Social media, Political Dissent, Online Experiment, China
JEL Classification: D72, L82, P16

*Department of Economics, Boston University. Email: rfisman@bu.edu.

†School of Finance, Renmin University of China. Email: yying.li@ruc.edu.cn.

‡School of Finance, Renmin University of China. Email: miaomeng@ruc.edu.cn. We are very grateful to David Stromberg for his valuable feedback and advice. This study received IRB approval from the Academic Ethics Committee of the School of Finance at Renmin University of China in September 2023. The experiment is registered in the AEA Randomized Controlled Trials registry under ID AEARCTR-0015231, available at: https://doi.org/10.1257/rct.15231-1.0

1 Introduction

In both electoral democracies and autocratic regimes, political leaders and aspirants vie for the support of citizens through direct outreach and political advertising. This has led to an extensive and growing literature on political persuasion that has focused in large part on television ad campaigns (Fowler *et al.*, 2021). However, the channels of persuasion have shifted online, accompanied by concerns over the influence that even brief video clips or pop-up ads can have over viewers' attitudes and beliefs.

In this paper, we study the effectiveness of online persuasion via a field experiment that emulates as closely as possible the actual consumption of internet content, and employs as outcomes real-stakes decisions to provide feedback on government policy. The specific context is a controversial set of amendments to the Public Security Administration Punishments Law proposed by China's parliament in August 2023, to criminalize vaguely defined actions that were "harming national sentiments." The legal change was subject to a public consultation period, during which time citizens could provide feedback via an online portal. In the 30 days the law was subject to public feedback in September 2023, approximately 100,000 individuals submitted more than 120,000 comments, an indication of the controversy surrounding the proposed changes.[1]

Coincident with the comment period, the social media app TikTok (known as Douyin in China) and other platforms were flooded with videos supporting the proposed changes, stressing the importance of defending national sentiment. Many of these seemingly independent videos repeated strikingly similar rhetoric.[2]

To evaluate how such videos influence public attitudes, we conducted an online experiment during the law's public consultation period, through a survey administered to over 2,000 participants from across China. Each subject assigned to one of the treatment groups watched one or two representative one-minute videos sourced from TikTok

[1] The median number of commenters on government rules was 322 as of 2024. There is a long right tail, however – the mean number of commenters is 5,400; for example, just over 111,000 commenters provided feedback on major changes to the civil code that were proposed in 2018.

[2] The uniformity of language led to speculation that the videos shared a common source, though there is no direct evidence of this.

(Douyin). These videos fell into two broad categories: the first category included *nationalist* clips, which supported the new legal provisions by emphasizing the need to protect national sentiment; the second included *critical* clips, which encouraged participants to take a stand against the amendments by highlighting concerns about vague language and the potential for unchecked administrative power. For both nationalist and critical videos, participants saw either a "calm" version or a more "emotional" one. In all cases, the TikTok clips were sourced from materials that were widely circulating on the Internet at the time. Finally, a set of control participants responded to an otherwise identical survey but did not watch any video.[3] After viewing the videos, participants were prompted to express their opinions on the legislation–either in support of or in opposition to the amendments–via the survey, with the assurance that their responses would be submitted anonymously to the government feedback website.

We find that even these brief exposures significantly shifted respondents' attitudes and their willingness to voice dissent. Participants treated only with a critical clip were 6.6 percentage points more likely to submit comments suggesting revisions to the proposed amendments, a 58% increase relative to the control group base rate of 11.4%. Participants in the nationalist clip treatments were 4.8 percentage points less likely to submit oppositional feedback relative to the control group (a 42% reduction) , and 4.0 percentage points more likely to write in support of the amendments (a 71% increase relative to the control base rate of 5.6%); we observe no effect of the critical clip on supportive comments). Interestingly, the latter effects were more pronounced for exposure to the emotional rather than the calm version and, somewhat surprisingly (and counter to our priors), the effect of the critical clip was stronger for the calm version.

These large effects raise the question of why even brief messages are so persuasive. The most immediate explanation is that they provide new information, leading viewers to update their beliefs about the proposed legal changes (and possibly this information is more salient when the message is delivered in a dramatic and emotional manner). Alternatively, the videos may influence viewers' beliefs about the extent of support for

[3] We discuss in Section 3.1 the potential ethical concerns raised by treating citizens with exposure to these messages.

the amendments – and relatedly, how others might view dissenters. If these beliefs in turn influence participants' willingness to express a point of view, the nationalist clip would reduce criticism of what is perceived as a popular policy (and by the same rationale, the critical clip would increase willingness to provide feedback). Finally, the clips may affect willingness to provide feedback due to demand effects, to the extent that they affect what respondents see as the "right" answer, from either the experimenters' or government's perspective.

A further set of analyses suggest that our main results may be driven largely by changes in second-order beliefs, i.e., the effect of videos on beliefs about perceptions of how others perceive the law and those who voice dissent. We first note that in our original experiment, before viewing the video, participants rated their own tolerance of opposition to the law to be higher than the tolerance that they perceived for others.[4] Second, while not obtained through experimental variation, we find that perceived intolerance of others is a very strong predictor of whether a respondent expresses support for or opposition to the law – as strong a predictor of own-intolerance.

Building on these insights, we conducted a supplementary experiment 7-10 days after our initial survey (still within the 30-day window during which the law was open to citizen feedback), to probe the extent to which misperceptions about the tolerance of others causally affects willingness to provide feedback. This supplementary experiment differed in two ways from the original one. First, given the smaller sample size (just over 500 respondents, compared to more than 2,000 for the main study), we focused only on the nationalist treatments, exploiting experimental variation between those viewing the calm versus emotional versions. Second, after viewing the experiment, respondents were given information about how past participants (i.e., those in the original experiment) responded to the question on how much expressing oppositional views of the proposed legal changes reflected a lack of patriotism. In contrast to the main experiment, we find that the emotional treatment is no longer predictive of writing feedback – whether supportive or opposing – on the new law.

[4] On a 1-10 Likert Scale, participants rated their own belief that opposition to the law reflected "unpatriotic" as 5.1 on average, but estimated others' views as 5.7.

The preceding results also help, at least indirectly, to address the concern that the videos signal *government* opinion on the law, and as a result may affect participants' beliefs about government retaliation (Madan *et al.*, 2023). More broadly, while we cannot fully rule out a role for social desirability bias, we took steps to minimize such concerns (some of which were also necessary to minimize any actual risk to study participants). These included removing content from the nationalist clips that might, in participants' minds, link it to government endorsement, and assuring participants that no identifying information would be collected.[5] We also explored the possibility that government-related concerns played a role in driving participants' responses via a supplementary experiment with 312 new participants. We asked these respondents, via a list design, to assess public perceptions of the extent to which the government would gain from the proposed legal provisions and, more importantly, whether they thought individuals would be targeted for expressing dissenting views. We find that neither concern was significant. [6]

We believe that our paper is the first to capture directly the consequences of short social media clips on political preferences or beliefs. We view our targeted study as complementing large-scale experiments that alter the social media feeds of study participants. An ambitious pair of studies focused on Facebook and Instagram users during the 2020 U.S. presidential election campaign documents relatively muted impacts of limiting political content (Allcott *et al.*, 2024) or altering the algorithm governing platforms' feeds (Guess *et al.*, 2023). Our results suggest at least the possibility of a short-term impact

[5] Specifically, we removed the opening segments of some short videos, as they might be mistakenly perceived as coming from officially endorsed channels.

[6] Because we are also interested in the longer-term effect of the brief videos we study here, particularly the nationalist ones, we conducted follow-up surveys at 6 months and 12 months after the baseline experiment, collecting a total of 333 valid responses, a response rate of 16%. In the follow-up experiments, we presented the same information as in the baseline, including the legal provision, and asked respondents to report their intolerance levels (both self-reported and predicted for others) along with their attitude toward the legal provision. (Participants were not asked to watch the video again, since we are interested in the longer-term effects of the initial viewing, and their willingness to provide feedback was purely hypothetical, as the comment period had long since ended.) It is perhaps unsurprising that after 6 months, most respondents had little or no recall of the content of the proposed legal provision (90%) nor the videos they had watched (82%). Furthermore, the effect of the videos was no longer detectable in stated willingness to provide feedback. Yet from the perspective of influencing short-term beliefs and preferences—which is more relevant for impacting responsiveness during the brief comment window—our main results suggest a substantial impact. It is also outside the scope of our experimental design to assess the impact of repeated exposure to short video clips that carry a similar message, as would be the "treatment" for many TikTok users.

in a somewhat distinct political context, and using a design that allows us to explore potential channels through which social media exposure can impact political behavior.

Our findings on the role of first- versus second-order beliefs shed light on the mechanisms through which social media exposure can impact preferences. Given the role we find for the social intolerance of others, our work contributes to the growing body of research on how perceptions and misperceptions of others can reinforce conformity to a set of social expectations and/or willingness to show disapproval of establishment opinions (Cantoni *et al.*, 2019; Bursztyn *et al.*, 2020). Given our focus on (biased) second-order beliefs, our results also show the potential relevance of the spiral of silence (the unwillingness to express opinions that are potentially erroneously believed to be contrary to social consensus; see Noelle-Neumann 1974) for political discourse. On this theme, our work complements the findings of Huang & Ho (2024), which explores how the spiral of silence can be disrupted via an information intervention in a lab setting involving Berkeley undergraduates. While "spiral of silence" is often seen as a social phenomenon that may be absent in our context, past work has shown that even in anonymous settings, social norms can have decisive effects on individual behavior, and the influential work of Postmes *et al.* (2001) finds that social norms can be even more impactful in anonymous settings.

The particular features of our experiment allow us to consider the interaction between social media exposure and the spiral of silence. In addition, we are able to do so in a setting in which there may be legitimate fear of reprisal for dissenting viewpoints, and in a collectivist society in which concerns of social isolation may be stronger (Postmes *et al.*, 2001; Huang, 2005; Gearhart & Zhang, 2018; Bursztyn & Yang, 2022).

2 Institutional Background

2.1 Public Opinion Feedback in the Chinese Legislative Process

China's central government initiated online public participation in its legislative process in July 2005 when the draft Property Law was published by the *Legislative Affairs*

Commission of the National People's Congress (NPC) Standing Committee on the NPC official website for public input.

The feedback period typically extends for one month, allowing citizens to submit anonymous feedback through the NPC website.[7] Individuals can offer multiple comments, each potentially addressing specific clauses of the draft. While the total comment count as well as the total number of commenters is visible, individual comments remain private unless commenters opt to publish their feedback on other platforms.[8] Once the feedback period ends, only the draft and the number of submitted comments are displayed.

In April 2008, the NPC Standing Committee's Meeting institutionalized public participation by mandating that all draft laws be published online to collect public input. Since then, it has become standard practice to invite public opinion on proposed legislation. As of the end of 2024 the NPC website had recorded over 300 public consultations, many of which involved multiple revisions. Laws that closely impact daily life, such as the Labor Contract Law, the Law of the People's Republic of China on the Protection of Women's and Children's Rights and Interests, and the Civil Code, have drawn significant public engagement. For example, the 2018 revisions to the Civil Code attracted 440,491 comments from 111,208 participants (as compared to the median number of comments and commenters, which are 949 and 322 respectively).[9]

After the collection of public feedback, the legislative process requires an explanation of which clauses were amended based on public consultation, especially those with substantial comments.[10]

---

[7] See Balla *et al.* (2025) for a more detailed discussion of the Notice and Comment process in China, which bears some similarity to the U.S. process. In short, Chinese citizens can submit their opinions on proposed amendments to major laws, such as the Public Security Law we study here, via the NPC website (www.npc.gov.cn), while other laws may be handled by specific agencies that oversee them.

[8] Particularly for controversial amendments, some interested citizens and legal experts may choose to post their submitted comments on social media sites like Weibo and TikTok to signal their stance and rally similar feedback from the public.

[9] The *Civil Code of the People' s Republic of China* is the first law named as a code in New China, holding a foundational position in the legal system, passed in 2020. It serves as a declaration and guarantee of civil rights. If the Constitution focuses on restricting public power, the *Civil Code* emphasizes protecting private rights. Almost all civil activities, from major concerns, such as signing contracts and establishing companies to minor ones, such as paying property fees or obtaining a divorce, can find their basis in this law.

[10] For the full text of Legislation Law of the People' s Republic of China in English translation, see http://en.npc.gov.cn.cdurl.cn/2023-03/13/c_938674.htm, last accessed July 18, 2025.

It also mandates justifications for any clauses that were commented on but left unchanged. In practice, however, the feedback details are rarely disclosed with the final legislation, leaving contributors unsure of the extent to which their suggestions were taken under consideration. Balla *et al.* (2025) show that rule-making is responsive to critical commentary, arguing that it serves as a means to gather information on potential dissent.

2.2 The Revision of the "Public Security Administration Punishment Law"

Within China's legal framework, the "Public Security Administration Punishments Law" (abbreviated to Public Security Law hereafter) is of considerable significance, covering a wide range of behaviors similar to those addressed by China's Criminal Law, but for less severe offenses. The Public Security Law is also quite recent, passed in 2005 (whereas the original Criminal Law was ratified in 1979). [11] In contrast to criminal statutes, the Public Security Law follows a simplified procedure that gives public security agencies broader powers to impose penalties, including detention of up to one month, without requiring involvement from a public prosecutor or the judiciary. If a penalty is later deemed erroneous, individuals can seek state compensation, but this only occurs after the payment of a fine or detention has already taken place.

The law includes vaguely defined offenses like "picking quarrels and provoking trouble," which enabled authorities to use broad charges to detain individuals for activities such as petitioning the high-level government authorities to address local injustices.[12]

In June 2021, 16 years after its initial enactment, the central government proposed revisions to the Public Security Law, with a Notice and Comment period of September 1-30. During this time, public feedback was collected on the NPC website. Key areas in the draft include amendments on information security and privacy rights, but the most widely discussed change was the addition of Article 34, which aimed to prevent actions

[11] For the primary text of the Public Security Law, see here: https://www.gov.cn/ziliao/flfg/2005-08/29/content_27130.htm; for Criminal Law, see here: http://www.npc.gov.cn/zgrdw/englishnpc/Law/2007-12/13/content_1384075.htm, last accessed July 18, 2025.

[12] Shangfang (上访), or petitioning, is a traditional mechanism in China whereby individuals seek redress by bringing grievances—often regarding perceived injustices by local authorities—to higher-level government bodies. Although legal in principle, petitioners are sometimes subject to surveillance, harassment, or detention, particularly when their actions are viewed as disruptive.

that were "disrespectful toward national heroes" or harmed national sentiments. The more controversial elements included vaguely worded prohibitions against: (1) "wearing clothing or symbols in public places that are detrimental to the spirit of the Chinese people and harming national sentiments of Chinese people, or forcing others to do so;" and (2) "producing, transmitting, promoting, or disseminating items or speech that is detrimental to the spirit of the Chinese people and harming national sentiments of Chinese people." Violations could lead to detention of five to ten days and fines of up to 5,000 RMB (or just under US$700).

This amendment sparked widespread debate online. Legal professionals cautioned that phrases like "detrimental to the spirit of the Chinese people" were too abstract and vague, leaving them open to subjective interpretation by law enforcement. This ambiguity, they warned, could result in excessive discretion and even abuses of power. At the same time, pro-nationalist videos endorsing these terms were widely circulated. Supporters of these videos label critics as anti-government, accusing them of sowing distrust in law enforcement. Anecdotally, it appeared that videos endorsing Article 34 far outnumbered those critical of it, so that if they were persuasive to potential critics, the videos could blunt dissent.

3 Experimental Design

From September 1 to September 30, 2023, the NPC solicited public feedback on the draft amendments to the Public Security Law through its official website. We conducted our experiment during this period on WenJuanXing, China's largest professional online survey platform.[13]

There were three phases to the experiment: a pilot; the main experiment; and a supplementary module that aimed to allow us to refine the interpretation of our main results. The three involved a total of 3,342 participants, and generated a total of 3,083 valid questionnaires (113 in the pilot; 2122 in the main experiment; and 848 in two

[13] WenJuanXing is the largest market survey company in China. While it has not been extensively used by economics researchers, it has had over 100 Global Fortune 500 companies as clients.

supplementary experiments). An overview of the various parts to each of the three phases can be seen in Figure 1.

We begin by describing the main experiment, which took place September 10-20, 2023. Respondents were randomly assigned to one of four groups, and within the first two, they were further randomly assigned to two subgroups, as described below.[14]

In all but the control group, respondents were shown videos related to the proposed changes to the Public Security Law. Note that in all cases, the videos were re-edited to mask the publisher's identity to reduce the chances that they might be associated with government-affiliated entities.

Nationalist Treatment: Respondents were shown one of two videos that supported the proposed legal changes, presenting the amendments as a matter of nationalist, and thus discouraging opposition. We selected the videos from among those with over 300,000 "likes" on TikTok, to ensure that the clips selected represented content that TikTok users would plausibly encounter in reality. Approximately half (52.8%) of respondents were shown a clip that featured an emotional narrator who spoke disparagingly of opponents to the legal change. We refer to this video as the *emotional nationalist* clip. The other half of respondents were randomly assigned to a *calm nationalist* clip, in which a speaker in a more measured tone rebuts concerns expressed by those opposed to the legal change.

Critical Treatments: Respondents in this group watched videos that critiqued the arbitrariness of the Public Security Law and the lack of regard for legal principle. The videos also described the proposed provisions as intentionally vague in order to permit broad discretionary power, thus raising concerns about government overreach and potential abuse of authority. Again, we used one video that used more forceful rhetoric to criticize the discretion that the new law would give to officials (*emotional critical* clip), while the other took a more muted tone, explaining how vague legal provisions could leave enforcement guidelines unclear (*calm critical* clip). As with the nationalist treatments, approximately half were assigned to each of the emotional and critical subgroups.

[14] The detailed steps of the main experiment can be seen in Appendix Figure A1. To randomize participants, we presented them with a set of 6 randomly ordered characters (À, Á, Â, Ã, Ä, and Å). Respondents picked a character, which determined their assignment.

Mixed Treatments: Participants in this group watched both a nationalist video and a critical video (again both emotional), with the order of presentation randomized. Since we find no difference throughout on ordering, we simply combine all mixed treatment respondents into a single group.

Control Group: Participants in this group did not watch any related videos.

This set of treatments allowed us to analyze the differential impacts of nationalist and critical video content, as well as the combined effect of both, on respondents' attitudes and behaviors toward the proposed legal provisions. Appendix Table A1 outlines the themes and representative phrases from each video used in the various treatments.

Immediately following the treatment, participants were asked about their willingness to provide an opinion on the new law, which is our first and most direct measure of the treatments' impacts. Specifically, we ask: "Do you have any opinions on this newly added regulation? (您对此新增条例有无修改意见?)" While in English translation, the question is quite ambiguous about whether it is positive or negative, "opinion" (意见) tends to reflect oppositional views in Chinese. Consistent with this interpretation, responses to this question were very strongly correlated with ultimately providing oppositional feedback on the law ($\rho = 0.27$), while responses were completely uncorrelated with providing supportive feedback.

Respondents were then presented with several anecdotes that highlighted the potential impact of commenting. This part of the design was motivated by results from our pilot study, which involved 113 participants (see Figure 1). After watching a video (randomly selected from one of the four in the main experiment) pilot respondents were asked, "What do you think are the reasons why some people are unwilling to voice their opinions?" They could choose any number of responses from a list of seven options.[15] The three primary reasons we identified were as follows: 1) They felt that as long as they complied with the law, the expansion of administrative power would not affect them

[15] The choices included: (1) This regulation doesn' t concern me -—it will not affect my life; (2) It's too much trouble – I can't be bothered; (3) I am not an expert so I cannot think of better ways of writing the regulation; (4) I support the regulation so there is no need for changes; (5) With so many people in our country, it' s not my role to give feedback.; (6) Writing feedback is pointless – the government will not listen to us anyway; (7) There might not be much opposition so my opinion would go against the mainstream.

personally, 2) They could free-ride on the efforts of others, and 3) They believed their opinions would be ignored.

We provided all participants with information that addressed these potential concerns about the benefits of providing feedback, via a set of three examples that illustrated: 1) the potential for misuse of public power; 2) the public's general enthusiasm for providing feedback on proposed laws; and 3) the possibility of changing legislation via public campaign. We describe the specific illustrative examples in more detail in Section 5 and the full text is available in Appendix A.

After each illustrative example was provided, respondents were again queried about their willingness to provide feedback on the new law, to track how the treatment effect interacted with knowledge about feedback efficacy.

The main experiment ended with an invitation to write down specific suggestions for the proposed changes. We informed respondents that their comments would remain anonymous, and that we would submit their suggestions to the official website (anonymously) on their behalf. Each comment was submitted separately through the text box provided on the government website, without any attribution or mention of the comment's author, and respondents were informed in the questionnaire that their responses would be anonymized and uploaded to the website by the research team, with no additional information provided.[16] If they did not have any suggestions or comments, they could leave the response blank. Finally, we collected basic demographic information, including their hometown, age, gender, education, income, occupation, and attitudes toward money, trust, and cooperation.

We took a number of steps to ensure the reliability of responses. First, the particular setting affords at least some advantages relative to non-China data collection, in that there is a more rigorous linkage of respondents to actual individuals. The WenJuanXing platform requires mobile phone verification for registration, which is again authenticated when an individual joins the sample. The redemption of any payment requires cross-verification via real names and government identification, and the name associated with

[16] Apart from ensuring the safety of our subjects, we would be unable to verify or record any comments made directly by participants, since all feedback on the website is confidential.

the bank card linked to WeChat. We emphasize, however, that no identifying information was retained – as we discuss below in Section 3.1, at our request, the company did not keep the data from our experiment, and we only received de-identified data.

At several points during the survey, simple questions were included to screen out bots and inattentive respondents. For example: "Please enter the number 43," and "12+21 =?" and, "What is the capital of China? Please choose the word below that appears in the preceding sentence." with the options Capital, Beijing, Shanghai, and Nanjing. A less attentive reader would choose Beijing, whereas Capital is the correct response.

We further excluded responses from the video-watching group that were completed in under 3 minutes, as recorded by the survey platform (5.3 minutes for the control group), and responses for which participants consistently selected the same option across all questions (e.g., always choosing the first option).

Our baseline experiment began with a sample size of 2,206, resulting in 2,122 valid responses after exclusions.

To refine our interpretation of the results from the main experiment, we conducted two supplementary surveys during September 21-30, 2023. In Supplementary Experiment 1, we surveyed 536 additional respondents who were all shown one nationalist video. We took this approach for several reasons. Most importantly, given the limited resources to conduct the follow-up, we wanted to limit the experiment to two treatments. And in doing so, we wanted to maintain as much comparability between treatments as possible, which is why we elected to have the calm version serve as a benchmark, given that – as we will see below – it showed relatively modest treatment effects in our main experiment. In this follow-up experiment, after viewing the video, all participants were then provided with information on *others'* beliefs about the new law based on responses of participants in the main experiment to the question, "Do you think that people who propose amendments to the new legal provision are unpatriotic?" Specifically, we informed participants in this supplementary survey that, "The average score for the 2,212 respondents who previously participated in this survey, for the question "Do you think that people who propose amendments to the new legal provision are unpatriotic?"is 5.14 (below the midpoint

of 5.5 on a scale from 1 to 10, where 1 represents "I do not think so at all" and 10 represents "They are very unpatriotic." The purpose of this supplement was to assess the role of second-order beliefs in driving respondents' willingness to provide feedback on the proposed law.

In a second supplementary survey, we conducted two list experiments with 312 new respondents, who all watched the emotional nationalist video, and were then asked to assess their attitudes toward the government and its potential responses to oppositional feedback from citizens. Of particular note, among the items in the list were "The government can benefit from this revision," and "The government will punish citizens who voice opposing opinions."

Finally, we conducted follow-ups with a subset of respondents from the baseline experiment six and twelve months after the initial survey; however, as we have already noted in the introduction, the response rates (and recall) for these follow-ups were quite low.

3.1 Discussion of Experimental Ethics

We took very seriously protecting the interests of our subjects in designing and implementing our experiment, which involved considerations that were complementary to our own interests in obtaining valid responses from them.[17]

We identified two primary risks: first and foremost, exposure to direct retaliation from the government from participating in the experiment; and "treating" some subjects in a way that could permanently alter their beliefs – either positively or negatively – toward the government and its policies.

Our approach to managing the first risk was to ensure that all feedback that we submitted to the government website – and indeed all responses to the experiment – were done with complete anonymity. As noted earlier, this was also essential to persuading respondents to comment critically on the proposed law. To eliminate even the residual risk

[17] The experiment was conducted with fast-tracked permission from Renmin University's Academic Commission, the institution's equivalent of an IRB. The third author was not involved with the project until after all data had been collected and given, as we describe below, that all data are completely anonymized, received an IRB exemption.

that subjects' identities might be uncovered later, after the survey was fully completed, we instructed the survey company to destroy the data.[18] At the time of writing, there existed no way of linking survey answers to individual respondents.

The second risk we identified was that participation in the experiment could alter participants' views on the pending Public Security Law and the government in general. We saw this risk as minimal ex ante, as the videos we used in our treatment were commonplace on the Internet, and hence represented exposure to material that they might encounter in everyday life. Relatedly, ex post we found that the consequence of exposure was relatively short-lived – while response rates for our longer-term follow-up were limited, among respondents we were able to contact, we found no relationship between exposure to videos – whether nationalist or critical – and attitudes toward the government and/or the Public Security Law.

3.2 Variable Construction

Our main analysis is built on whether and how a respondent chooses to share an opinion on the Public Security Law. This opinion can be supportive, oppositional, or (less commonly) neutral. Of our sample of 2122 participants, 494 (23.3%) wrote at least something in the text box we provided for feedback via the government web portal. To determine whether a comment was positive or negative, we had three research assistants classify each as supportive, oppositional, or neutral. In the vast majority of cases (99.25%), they were in uniform agreement. In all such cases, we follow the consensus classification. In the 16 instances in which the RAs disagreed, we classify the comment as neutral (though we found that we generally agreed with the majority view; and as expected, the classification of these 16 observations does not affect any of our results).

Of the 494 comments, 90 were classified as neutral. These tended to be very short comments (median length of just 6 characters, in practice usually just one or two English words. Casual inspection reveals that these were almost always some variant on "no"

[18] It would be an unprecedented step for the government to try to access data from WenJuanXing. The company itself has no ties to the Chinese government and is in fact majority-owned by a Hong Kong conglomerate that is, in turn, held by a diverse set of international investors.

(as in I have no comments) or "no opinion" or, for the longer ones, "I have no relevant expertise to comment on this law." In practice, we thus treat these cases as equivalent to non-responses.

Of the 404 remaining comments, 265 (65.6%) were classified as opposing of the new law, while the remaining 139 (34.4%) were classified as supportive. These varied enormously in length, including several one-word endorsements (e.g., "Good") and also some that were many hundreds of characters. In general, critical comments were longer: the median number of characters was 36 and 54 for supportive and opposing comments respectively. We define indicator variables, $\mathbb{1}(Oppose)$ and $\mathbb{1}(Support)$ that reflect these opposing and supportive comments.

The five covariates of primary interest are indicator variables for each treatment arm (*Control* is the omitted category: $Nat_Emotional$, Nat_Calm, $Crit_Emotional$, $Crit_Calm$, and $Mixed$).

Finally, we include basic sociodemographic controls and province fixed effects, to improve precision (assuming random assignment, these should not affect our point estimates, which is true in practice).

The controls include basic demographics (categorical variables for age, income, education, and gender); and control for prior knowledge of the proposed reform.

3.3 Sample Description

Table 1 presents descriptive statistics for the full sample, with balance tests shown in Appendix Table A2. The most notable finding is a significant difference in gender distribution between the $Nat_Emotional$ and *Control* groups.

We also map the provincial distribution of respondents, as shown in Figure A2 in Appendix A. The map shows that respondents are drawn from a very broad geography within China.

4 Results from the Main Experiment

We begin by presenting our main results in graphical form in Figure 2, which summarizes the percentage of respondents in each group that provided supportive or opposing comments. The dotted vertical lines provide the benchmark percentages from the control group; the dark bars to the left represent the proportion of opposing comments and the lighter bars to the right represent the proportion of supportive comments; we also include "whiskers" indicating the 95% confidence intervals for each fraction.

We begin at the top of the figure with the nationalist treatments, where we the endpoints of all bars are shifted to the right indicating fewer opposing comments and more supportive ones, relative to the control. This effect is more pronounced for the emotional version – in fact, there is no difference between the calm version and the control group for supportive comments.

We observe the opposite effect for critical treatments, though the effect is much more pronounced for opposing comments – there is essentially no effect of watching critical clips on willingness to provide supportive comments. Curiously – and counter to our priors – in this case the effect is driven entirely by the calm, only-the-facts version.

Finally, we observe no treatment effect in the mixed arm, in which participants saw both nationalist and critical clips, suggesting that the two cancel one another out.

Our main results are also presented in Table 2 using a regression framework, which allows us to include fine-grained individual controls. Given that our randomization was largely successful, it is as expected that the patterns are very similar to those appearing in Figure 2. For each of $\mathbb{1}(Oppose)$ and $\mathbb{1}(Support)$, we present three specifications, all based on the following:

$$Outcome_i = \beta_n Nationalist_i + \beta_c Critical_i + \beta_m Mixed_i + X_i + \gamma_{p(i)} + \varepsilon_i \tag{1}$$

where i indexes the individual respondent; $Outcome$ is either $\mathbb{1}(Oppose)$ or $\mathbb{1}(Support)$; $Nationalist$, $Critical$, and $Mixed$ are indicator variables for the relevant treatments (or as appropriate a disaggregation of these into emotional and calm); X is a set of i

characteristics, and γ are fixed effects for i's province of residence $p(i)$. The results are consistent with what we observe in our graphical presentation of the results.

The regression coefficients provide straightforward estimates of the treatment effects. In column 2, we observe that for the nationalist clips, there is a 4.8 percentage point reduction in the likelihood of providing oppositional feedback,a 42% reduction relative to the control group mean of 11.4%; for critical videos, there is a 6.6 percentage increase in the probability of providing oppositional feedback (58%) (column 2). There is a much bigger effect from the emotional video for the nationalist treatments, while somewhat surprisingly, for critical treatments the calm video has a bigger effect. The nationalist treatments have a very large relative effect on the likelihood of supportive impact – the point estimate of 0.04 implies a 71% increase relative to the control group mean of 5.6%. (There is no impact of critical treatments on supportive comments.

As we explained in Section 3, after viewing the video, we provided several illustrative examples to convey to respondents the extent to which their comments could be impactful. This element to the experiment allowed us to examine how our treatments interact with potential shocks to empowerment (or disempowerment) – a topic we take up in Section 5 – but does raise the question of how video exposure would have impacted willingness to provide feedback in the absence of such messages. To provide a "pure" measure of the video treatments' impacts, we asked participants immediately after watching the video about their willingness to provide an opinion on the law (which, recall from the discussion in Section 3, connoted an *oppositional* opinion). We look at responses to this question, which was asked immediately followed the video(s), as an alternative outcome in Appendix Table A3. We observe patterns that are very closely aligned with our analyses that use $\mathbb{1}(Oppose)$ as the outcome – participants who saw nationalist videos, especially the emotional variant, were less likely to express a willingness to provide feedback; critical videos had the opposite effect.

4.1 The Role of Potential Government Retaliation

The experiment was designed to minimize the concerns of respondents about government retribution, but such "demand effects" remain a concern given the government's own interests in the success of the law.[19] This concern may have been further reinforced by the government's ability to trace individuals online through telephone numbers via "real-name registration" (though it is precisely this registration that gives us greater confidence in the quality of our sample). To the extent that respondents who viewed the nationalist videos as government-endorsed messages could prompt behavioral adjustments due to fear of retaliation.

As noted earlier, we made several adjustments within the experiment to mitigate this concern: we provided assurance of anonymity in any feedback we provided on the new law; and when embedding short-video content into the questionnaire, we removed the publishers' account information, preventing respondents from viewing it as government-endorsed.

To examine whether such fears are, in fact, a major concern in the first place, one of our follow-up surveys presented an additional 312 participants with a list experiment. After completing the baseline questionnaire, participants were asked to participate in two list experiments, designed to capture sensitivity to government retribution via indirect responses. The control and treatment groups were both presented with four descriptions irrelevant to the study, but the treatment group also received two further statements: "The government can benefit from this revision" and "The government will punish citizens who voice opposing opinions." The full list of items can be found in Table A4, with the relevant experimental items highlighted. Latent concerns were calculated by comparing the mean item counts between the treatment and control groups.

As illustrated in Figure A3, treatment respondents demonstrated minimal mean differences compared to the control group, with average disparities of a little over 0.1. T-tests yielded statistically non-significant results in both list experiments. With a 1-

[19] See, e.g., Chen & Yang (2019); Lu *et al.* (2024). It is worth noting also that the government is invested in minimizing dissent, and as such is motivated to elicit feedback and preempt any substantive concerns that arise. See, e.g., Fisman *et al.* (2021).

sided test, we can reject that the fraction of participants who agreed with each statement is more than 0.2 at the 5% significant level.

Overall, the results of this follow-up survey suggest that worries about government retribution were unlikely to be a first order concern in governing participants' decisions (and even less likely that it would have a large differential effect on those who saw particular videos).

4.2 Self- versus Social-intolerance

There are two primary explanations for the impact of exposure to videos on respondents' willingness to provide feedback on the proposed law. Most obviously, it could change an individual's beliefs about the law's merits. Alternatively, it may alter their beliefs about whether others support the law.

To gain some insight into whether own-beliefs and/or beliefs about others' support were affected, after viewing the video(s) (or at the same point in the survey for the control group), respondents were asked the following pair of questions:

- *Do you think that people who propose amendments to the new legal provision are unpatriotic? /*
- *Do you think that others might see people who propose amendments to the new legal provision as unpatriotic?*

For both questions, responses were elicited on a 1-10 scale, with higher values indicating greater intolerance toward proposing amendments to the Public Security Law. Specifically, for the first question, 1 was labeled as "I do not think so at all" and 10 as "They are very unpatriotic." For the second, 1 was labeled as, "I do not think others would care" and 10 as, "Others will believe they are very unpatriotic." [20]

While responses are not directly comparable, respondents generally gave higher values for the social intolerance question – a mean of 5.6 versus 5.1 for self-intolerance. We define $Self_Intolerance$ and $Social_Intolerance$ based on responses to these questions.

[20] The literal translations of the text is somewhat different and relates to "worship of foreigners." The text here best reflects, in our view, how the questions would be perceived by a Chinese reader.

To provide some preliminary evidence on the role of "self-intolerance" and "social-intolerance" (i.e., second-order belief) in driving willingness to provide supportive or oppositional feedback on the proposed law, we proceed in two stages. First, we repeat the analysis in Equation (1), using as the outcomes $Self_Intolerance$ and $Social_Intolerance$. Second, we look at how these two variables correlate with our two outcomes of primary interest, $\mathbb{1}(Oppose)$ and $\mathbb{1}(Support)$. Before proceeding to this pair of results, we first note that while responses to the two questions are, unsurprisingly, quite correlated, there appears to be some independent variation – the raw pairwise correlation is 0.46.

We start with the link from treatment status to self- and social-intolerance in Table 3. In the first pair of columns, the outcome is self-intolerance, while in the second pair it is social-intolerance. The results are directionally similar for both outcomes, suggesting that nationalist videos increase intolerance of opposition to the proposed law, while critical videos have the opposite effect.

In Table 4 we look at how self- and social-intolerance in turn are correlated with the ultimate outcomes of $\mathbb{1}(Oppose)$ and $\mathbb{1}(Support)$. Focusing first on $\mathbb{1}(Oppose)$, in columns 1 and 2 we include each of self- and social-intolerance separately as regressors; both are highly predictive of abstaining from writing opposing comments. When both are included in column 3), we find that, despite their modestly high colinearity, each remains a significant predictor of $\mathbb{1}(Oppose)$, although the coefficient on $Self_Intolerance$ drops by nearly 40%. Finally, we perform a sort of mediation analysis in column 4, in which we include the various video treatments as well as the mediating variables $Self_Intolerance$ and $Social_Intolerance$ as regressors. Most notably, the coefficients on $Self_Intolerance$ and especially on $Social_Intolerance$ are essentially unchanged while that of $Nat_Emotional$ is reduced by 43%. We observe less of a "mediation effect" in predicting $\mathbb{1}(Support)$ (although the critical video treatments were never predictive of $\mathbb{1}(Support)$ in the first place).

4.2.1 Supplementary Experiment on Second-Order Beliefs

Naturally, the empirical exercise above is only loosely suggestive. In an attempt to distinguish more decisively between self- and social-intolerance as predictors of willingness to provide feedback, we ran a supplementary experiment inspired by our initial set of results. As discussed in Section 3, the setup in this experiment is very similar to the pair of nationalist treatments in the main experiment (for reasons of sample size, we focused on just two treatments, as explained earlier), though with the following information provided to participants after viewing the video but before being given the opportunity to provide feedback on the proposed law:

> "The average score for the 2,212 respondents who previously participated in this survey, for the question "Do you think that people who propose amendments to the new legal provision are unpatriotic?" is 5.14 (below the midpoint of 5.5 on a scale from 1 to 10, where 1 represents "I do not think so at all" and 10 represents "They are very unpatriotic."

For the 536 participants in this experiment, we repeat our main analysis of whether treatment impacted willingness to provide feedback on the Public Security Law; these results appear in Table 5. In columns 1 and 2 we present a variant on Equation (1) as follows:

$$Outcome_i = \beta_n Nat_Emotional_i + X_i + \gamma_{p(i)} + \varepsilon_i \tag{2}$$

with $\mathbb{1}(Oppose)$ as the outcome. Whether we focus on the bivariate regression (column 1 or one with controls (column 2), the coefficient on $Nat_Emotional$ is very close to zero (though given the imprecision of these estimates we also cannot reject a large effect). This contrasts with the pattern we see in column 3, in which we present an analogous regression from the main experiment, focused on the comparable subsample of participants that viewed one of the nationalist videos. In this sample, we find a significantly negative coefficient on $Nat_Emotional$ ($p < 0.05$). We pool the two groups in column 4, and include an indicator variable, $Informed$, denoting participants in the later experiment

in which subjects were informed of the average responses of others. As expected given the noisiness of our estimates in columns 1 and 2, the coefficient on $Nat_Emotional \times Informed$ is not statistically significant. Interestingly, the coefficient on $Informed$ is significantly positive (and particularly if we do not include the interaction term, $p < 0.01$), indicating a greater willingness to provide oppositional feedback in the later experiment. This could be due of the effect of the information treatment, which generally informed participants that social intolerance was lower than anticipated; or it could simply reflect differences in the sample because the informed group was surveyed at a later date.

We repeat these analyses in columns 5–8 for $\mathbb{1}(Support)$ as the outcome. In columns 5 and 6, the coefficient on $Nat_Emotional$ is close to zero, while it is large and significantly positive for the $Informed = 0$ (Main-Expt) sample in column 7; as observed in column 8, the difference between the two is statistically significant.[21]

Overall, the results in this section suggest that the videos we presented to participants impacted their willingness to provide feedback – both supportive and oppositional – at least in part because of their effect on second-order beliefs. When we informed subjects of others' actual tolerance, the treatment effect disappeared.

5 Additional Channels and Interpretations

In this section, we consider some further interpretations or potential sources of influence on willingness to provide feedback. While we do not believe that these alternatives can fully explain the set of results we have presented thus far, they could potentially play a role in how survey participants responded to some subset of our outcome measures.

5.1 Could the Treatment Affect Beliefs about Others' Willingness to Comment?

In the previous section, we showed that willingness to provide feedback is no longer affected by videos when accompanied by accurate information on others' tolerance for

[21] We find that the direct effect of $Informed$ also indicates a lower propensity to provide supportive feedback, though in this case the effect is not statistically significant.

criticism of the law. We interpret this as reflecting the fact that willingness to provide feedback is very much affected by a form of social pressure. Yet information on others' tolerance for such criticism could also be simply because it affects the number of other people they believe will provide feedback. Since free-riding on the efforts of others was highlighted in our pilot survey as a reason for not providing feedback on the law, we also included the following question after the video treatment: "What do you think are the reasons why some people are unwilling to voice their opinions?" It was still be presented as a multiple-choice item, with one option being: "With so many people in our country, it's not my role to give feedback." This option serves as a behavioral measure for whether respondents perceive opinion-sharing as a free-riding opportunity (i.e., believing others will provide feedback, thus exempting themselves).

We find that responses to this question are completely uncorrelated with treatment status, suggesting that second-order beliefs are unlikely to act through their impact on expectations about others' willingness to provide feedback.

5.2 Could the Treatment Effect of Videos be Affected by Beliefs about Efficacy of Commenting?

A number of factors related to the efficacy of commenting could mediate the impact of the video treatments on willingness to provide feedback. There were three factors that emerged from our pilot as primary determinants of willingness to comment on the proposed law. First, understanding the potentially harmful effects of vaguely-defined government powers could increase the impact the critical videos. Second, understanding that many others are likely to comment exacerbates the free-rider problem and thus makes people less likely to provide feedback – whether supportive or oppositional. Finally, learning that the government may in fact respond to public feedback could make responses *more* likely.

To explore whether (and how) these factors interacted with our main treatment, after viewing the video(s), *all* respondents were then provided with three illustrative examples to highlight each of the three issues mentioned above. After each piece of information, we

asked respondents about their willingness to provide (oppositional) feedback on the new law.

Specifically, we informed participants of (1) the number of opinions others have submitted on previously proposed legislation, to reduce the perception that their input is unnecessary; (2) examples of instances in which innocent individuals were harmed or even died due to abuses of power, highlighting the potentially widespread and severe consequences of broadly defined legal authority; and (3) the successful example of public feedback altering proposed laws, from the Law of the People's Republic of China on the Protection of Women's and Children's Rights and Interests, to emphasize that the People's Congress may respond to public sentiment. (see Table A1 for the exact wording of what respondents were told.)

As noted earlier, before providing any of this information (but after watching the video(s)) we asked participants about their willingness to provide an opposing view of the law; they were then asked the same question about willingness to provide feedback after each of the three pieces of information.[22]

Our main interest is whether giving respondents more accurate information about the actual efficacy and value of providing feedback affects the link from our video treatments to stated willingness to comment on the law. To capture this possibility, we look at how treatment status predicts willingness to provide oppositional comments, using specifications that parallel that of Equation 1. The outcome variables we employ are: $Willingness_t$, which captures responses to the question about willingness to provide oppositional feedback at time t, where $t = 0, 1, 2, 3$ indexes the number of informational treatments a participant has experienced. Our overall observation on these patterns is that there are at most marginal mediating effects of information on the videos' effect on willingness to provide oppositional feedback – throughout, nationalist videos has a negative effect (though not always significant) while critical videos have a positive effect. On balance, we find limited evidence that such information plays a mediating role.

[22] There is a significant positive correlation between self-reported willingness and the actual submission of opposing opinions, but no robust correlation with supporting opinions, as expected; see Appendix Table A5.

We see this finding as underscoring the relevance of our main result – exposure to brief social media content impacts willingness to express dissent, and is not reliant on "complementary" knowledge such as the efficacy of providing feedback.

5.3 Long-term Impacts

To explore possible longer-term impacts, we conducted follow-up surveys with participants in the main experiment 6 and 12 months after the initial study, obtaining 173 and 160 valid responses, respectively. Combining both longer-term surveys, the response rate was approximately 15%.

Perhaps unsurprisingly, there are no discernible long-term effects, with the results of the follow-up survey suggesting that participants had largely forgotten about both the proposed law as well as the videos they watched. As shown in Appendix Table A6, only 32 respondents (9.61%) in the follow-up surveys could clearly recall the specific content of the new provisions in the Public Security Law. Of those who watched the videos in our baseline experiment, approximately 61 respondents (18.32%) remembered the main content of the videos, and only 49 respondents (14.71%) expressed ongoing concern about developments related to the legal amendments.

Table A6 display respondents' perceptions of others' tolerance attitudes toward the law across treatment groups. Notably, the belief that expressing opposition is unpatriotic changed in all groups, generally converging toward the responses of the control group, which did not view the video. As shown in Table A7, few respondents in the follow-up survey maintained opinions similar to those in the initial experiment.

It would be surprising if we had found otherwise – our results are noteworthy even for their short-run impact, and it would stretch credulity to imagine that such an intervention would continue to affect beliefs after months or a year. Yet what may be relevant from the perspective of government persuasion is precisely the short-run impact we document, which mutes potential opposition at exactly the moment when dissent could coalesce. Indeed, our results perhaps mirror the fleeting nature of public interest more generally, so that the type of persuasion we study here can be deployed in the short-term to manage

public opinion.

5.4 What Might the Impact of TikTok Videos on the Public's Response to the Public Security Law?

Even a rough calculation of the overall impact on comments related to the new law is beyond the scope of what is feasible in the context of our paper. It would require, for example, taking a position on the cumulative rather than individual impacts of video exposure, and accounting for the more focused experience that respondents in our experiment may have had relative to the casual TikTok user. Further, we would need to account for the extent to which TikTok users are more (or less) likely to engage in civic discourse relative to the average Chinese citizen.

Yet we can provide some facts on exposure to the types of videos we deployed which indicate their potential for limiting dissenting views. During the comment period, we canvassed the set of videos with titles or slogans related to the proposed legal changes (e.g., keywords such as "Public Security Administration Punishments Law," "harming national sentiments," "criminalizing clothing," and so forth) on China's two most popular short video platforms TikTok and Kuaishou. We focused on videos receiving at least 1,000 "likes."[23] The 50 most-related videos returned by this search had approximately 27 million views during September 2023. But it was very heavily skewed toward nationalist videos – 30 of the 50 videos were supportive of the law, and nearly 90 percent of total views were of the nationalist videos, 81 percent of which we classified as "emotional" rather than calm (see Appendix Table A8).

Given these figures, combined with our findings that nationalist videos substantially reduced willingness to provide critical comment, we see the emergent technology of short video platforms as an effective means of controlling public dissent. While it is of course impossible to assess the link between the ubiquity of nationalist (versus critical) video content online, it is also notable that the Public Security Law – including the more am-

[23] Based on the financial reports from the third quarter of 2024, TikTok's daily active users exceeded 600 million in 2024, while Kuaishou's daily active user count stood at over 400 million. Also note that 1,000 likes is a relatively high number, given that anyone can upload videos on TikTok, yet few videos get much attention, and there is a very long right tail of video popularity.

biguous provisions that may expand government power – were not subject to substantial revision. In the revised version of the law, which was released in July 2024, some greater specificity was added, consistent with at least some government responsiveness to critical feedback. For example, "'harming national sentiments" was modified to read, "clothing or symbols that promote or glorify wars of aggression or acts of aggression, refusing to heed dissuasion, and causing adverse social impact." However, as noted by China Law Translate (2024), the revised clause still lacks precise definitions of prohibited actions, as it continues to use (legally) ill-defined terms such as "promote" and "glorify."

6 Conclusion

We conducted an experiment with well over 2,000 respondents to study how online content impacts willingness to express dissent or support for government policies. The particular setting is the comment period for the Public Security Law during September 2023. In the main part of our experiment, respondents were exposed to either nationalist or critical videos (in an emotional or calm variant). We find that particularly emotional nationalist videos have a substantial impact on willingness to provide supportive (a 8.1 percentage points increase) or oppositional (a 7.1 percentage points decrease) feedback on the law through a government portal designed for eliciting public comment. We find that this effect disappears when we provide information on the extent to which others disapprove of those who criticize the law. Overall, our findings suggest that online content impacts the public's willingness to criticize government policy, in large part because it shapes their views on how dissent might be viewed by others.

Our findings highlight the potentially important role that social media may play in shaping public policy discourse, a point emphasized by Treisman & Guriev (2023) among others. There are many directions for future work that can build on our findings. In our setting, for example, holding a dissenting view may have social costs (given the importance of conformity in Chinese culture) but also potential risks given the government's tight control over dissent. Understanding the relative roles of each of these factors in driving our results is a promising area for future work. Relatedly, social media is prevalent

worldwide, across a range of political and economic circumstances. It may be interesting in future work to explore if and how exposure to online content differentially affects respondents in, say, the U.S. relative to China.

References

Allcott, Hunt, Gentzkow, Matthew, Mason, Winter, Wilkins, Arjun, Barberá, Pablo, Brown, Taylor, Cisneros, Juan Carlos, Crespo-Tenorio, Adriana, Dimmery, Drew, Freelon, Deen, *et al.* 2024. The effects of Facebook and Instagram on the 2020 election: A deactivation experiment. *Proceedings of the National Academy of Sciences*, **121**(21), e2321584121.

Balla, Steven J, Wan, Andrew, Xie, Zhoudan, Yeung, Yat To, & Zhai, Geng. 2025. Notice the comment? Chinese government responsiveness to public participation in the policymaking process. *Governance*, **38**(2), e12901.

Bursztyn, Leonardo, & Yang, David Y. 2022. Misperceptions about others. *Annual Review of Economics*, **14**(1), 425–452.

Bursztyn, Leonardo, Egorov, Georgy, & Fiorin, Stefano. 2020. From extreme to mainstream: The erosion of social norms. *American Economic Review*, **110**(11), 3522–3548.

Cantoni, Davide, Yang, David Y, Yuchtman, Noam, & Zhang, Y Jane. 2019. Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement. *The Quarterly Journal of Economics*, **134**(2), 1021–1077.

Chen, Yuyu, & Yang, David Y. 2019. The impact of media censorship: 1984 or brave new world? *American Economic Review*, **109**(6), 2294–2332.

China Law Translate. 2024. *English-Language Comparison Table on Public Security Administration Punishments Law*. Accessed: 2025-06-28.

Fisman, Raymond, Lin, Hui, Sun, Cong, Wang, Yongxiang, & Zhao, Daxuan. 2021. What motivates non-democratic leadership: Evidence from COVID-19 reopenings in China. *Journal of Public Economics*, **196**, 104389.

Fowler, Erika Franklin, Franz, Michael, & Ridout, Travis. 2021. *Political advertising in the United States*. Routledge.

Gearhart, Sherice, & Zhang, Weiwu. 2018. Same spiral, different day? Testing the spiral of silence across issue types. *Communication Research*, **45**(1), 34–54.

Guess, Andrew M, Malhotra, Neil, Pan, Jennifer, Barberá, Pablo, Allcott, Hunt, Brown, Taylor, Crespo-Tenorio, Adriana, Dimmery, Drew, Freelon, Deen, Gentzkow, Matthew, *et al.* 2023. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, **381**(6656), 398–404.

Huang, Huiping. 2005. A cross-cultural test of the spiral of silence. *International Journal of Public Opinion Research*, **17**(3), 324–345.

Huang, Yihong, & Ho, Yuen. 2024. Breaking the Spiral of Silence.

Inglehart, Ronald, Haerpfer, Christian, Moreno, Alejandro, Welzel, Christian, Kizilova, Kseniya, Diez-Medrano, Jaime, Lagos, Marta, Norris, Pippa, Ponarin, Eduard, & Puranen, Bi. 2022. *World Values Survey Round Seven (WVS-7) Country-Pooled Datafile*.

Lu, Yingdan, Schaefer, Jack, Park, Kunwoo, Joo, Jungseock, & Pan, Jennifer. 2024. How information flows from the world to China. *The International Journal of Press/Politics*, **29**(2), 305–327.

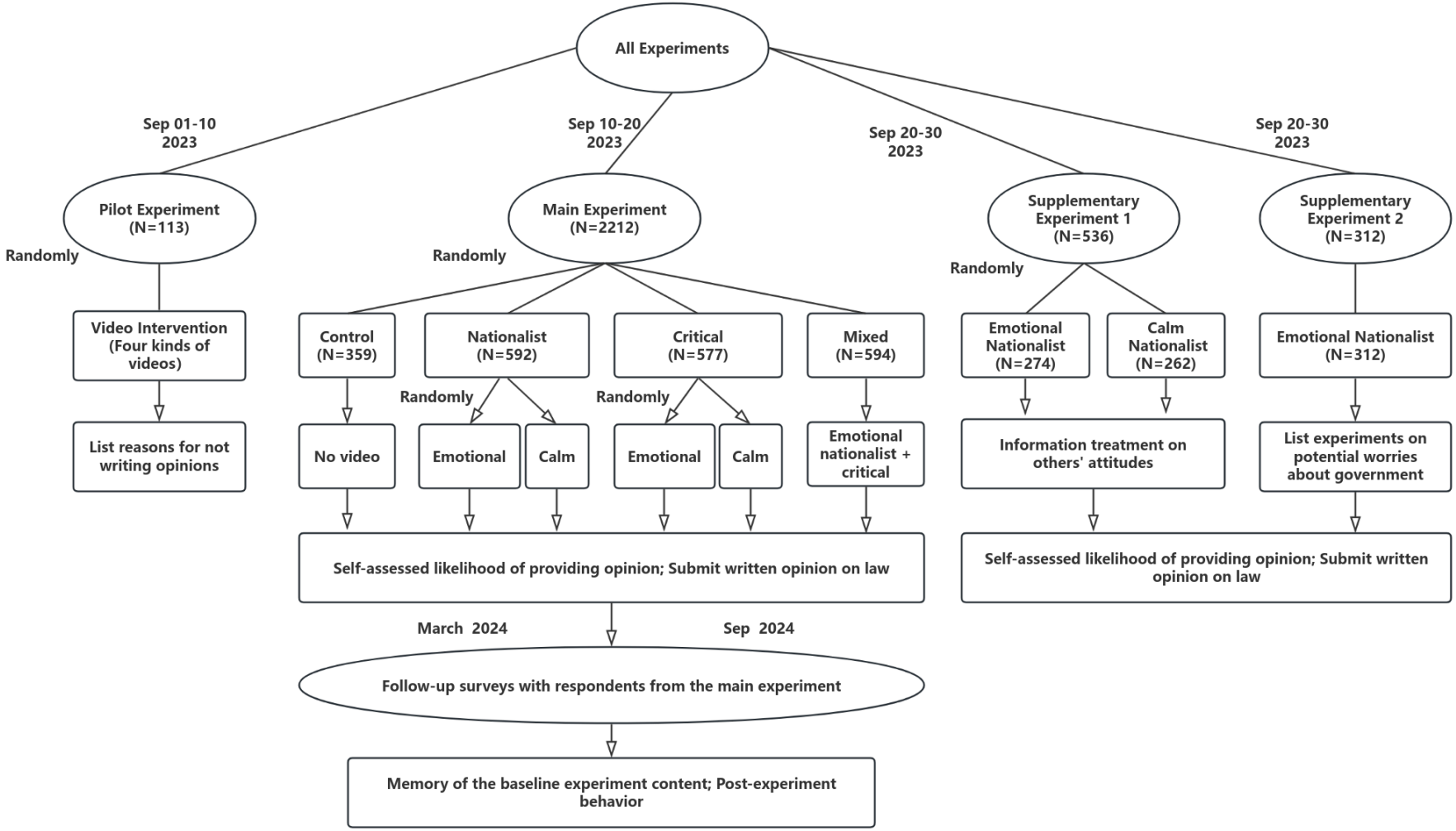
Madan, Shilpa, Savani, Krishna, & Katsikeas, Constantine S. 2023. Privacy please: Power distance and people' s responses to data breaches across countries. *Journal of International Business Studies*, **54**(4), 731–754.

Noelle-Neumann, Elisabeth. 1974. The spiral of silence a theory of public opinion. *Journal of communication*, **24**(2), 43–51.

Postmes, Tom, Spears, Russell, Sakhel, Khaled, & De Groot, Daphne. 2001. Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, **27**(10), 1243–1254.

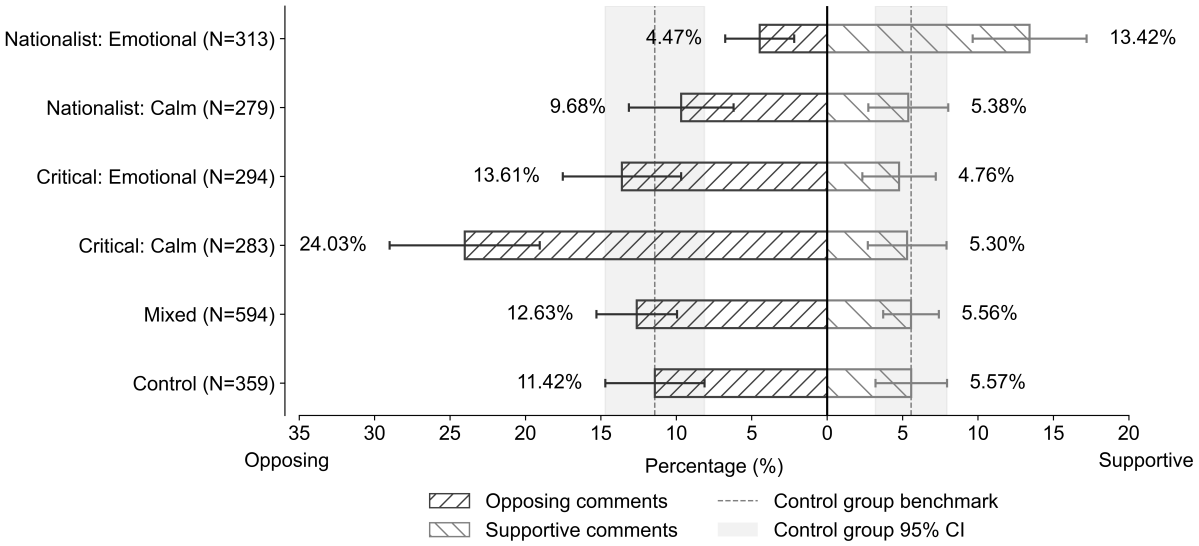
Treisman, Daniel, & Guriev, Sergei. 2023. Spin dictators: The changing face of tyranny in the 21st century.

Figure 1: Experimental Flow



Notes: This figure summarizes the design of our experiments. In both the baseline and supplementary experiments, participants watched one-minute videos. Before the main experiment, a pilot study was conducted to identify the primary three reasons why respondents were reluctant to express their opinions on the proposed Public Security Law. In both the baseline and supplementary experiments, participants were given targeted textual information addressing the top three reasons for withholding opinions. In supplementary experiment 1, new participants were treated with extra information on others' views on whether providing a oppositional feedback on the new law was "unpatriotic." In supplementary experiment 2, respondents participated in two list experiments. Follow-up surveys were administered six months and one year after the main experiment, repeating the same questions as in the main experiment, without any video or textual interventions.

Figure 2: The Likelihood of Commenting Across Different Groups



Notes: This figure shows the proportion of respondents that wrote opposing and supportive comments on the proposed provisions of the Public Security Administration Punishment Law. The proportion of opposing comments is the percentage of each group that suggested modifications or expressed dissatisfaction with the newly added provisions. The proportion of supportive comments is the percentage that wrote supportive comments and/or criticized those who opposed it. We also label the control group mean and 95% confidence intervals for ease of comparison.

Table 1: Summary Statistics for the Main Experiment

	Mean	SE	Min	Max
Sociodemographic Characteristics				
Gender	0.56	0.01	0	1
Age	3.40	0.02	1	6
Education	2.83	0.01	1	4
Income (yearly)	2.76	0.02	1	4
Attitudal Responses				
Patriotism	1.03	0.01	0	2
Cooperation	5.75	0.06	1	10
Money	7.65	0.04	1	10
Effort	7.51	0.04	1	10
Mistrust	6.40	0.04	1	10
Writing Performance				
Supportive Comments	0.07	0.25	0	1
Opposing Comments	0.12	0.33	0	1
Obs	2,122			

Notes: This table reports summary statistics for respondent characteristics in the main experiment. Variable definitions: *Gender* (0 = Male, 1 = Female); *Age* (1 = Below 18 years old, 2 = 19–25, 3 = 26–30, 4 = 30–40, 5 = 40–50, 6 = Above 50); *Education* (1 = High school and below, 2 = Associate degree, 3 = Bachelor's degree, 4 = Master's degree and above); *Income* (1 = Below 50,000 RMB, 2 = 50,000–130,000 RMB, 3 = 130,000–200,000 RMB, 4 = Above 200,000 RMB). In the "attitudal responses" section, "patriotism" is measured prior to any intervention (scored 0-2, measured by asking respondents which region in the world they would most like to settle in, 2= only chose China and does not want to live anywhere else. 1= first choice is China, but also selected other countries or regions. 0 = China is not among the regions where the respondent wants to live.). The other four items use questions based on Inglehart *et al.* (2022), involving beliefs about cooperation, money, effort, and trust, scored on a 1-10 scale. When respondents in the experiment provide supportive or opposing comments on the proposed regulations, the *Supportive Comments* or *Opposing Comments* are assigned a value of 1, respectively.

Table 2: Main Results of Video Intervention on Providing Comments

	𝟙[Oppose]			𝟙[Support]		
	(1)	(2)	(3)	(4)	(5)	(6)
Nationalist	-0.045**	-0.048**		0.041**	0.040**	
	(-2.272)	(-2.377)		(2.366)	(2.277)	
Nat_Emotional			-0.071***			0.081***
			(-3.302)			(3.525)
Nat_Calm			-0.023			-0.005
			(-0.946)			(-0.262)
Critical	0.073***	0.066***		-0.005	-0.004	
	(3.121)	(2.843)		(-0.360)	(-0.290)	
Crit_Emotional			0.020			-0.006
			(0.746)			(-0.332)
Crit_Calm			0.114***			-0.002
			(3.788)			(-0.100)
Mixed	0.012	0.003	0.003	-0.000	0.002	0.002
	(0.557)	(0.152)	(0.157)	(-0.010)	(0.114)	(0.154)
Feedback Rates (Control Group)	11.42%	11.42%	11.42%	5.57%	5.57%	5.57%
Controls	No	Yes	Yes	No	Yes	Yes
Province FE	No	Yes	Yes	No	Yes	Yes
Obs	2,122	2,121	2,121	2,122	2,121	2,121
R-squared	0.058	0.065	0.072	0.019	0.026	0.034

Notes: This table presents the baseline results of the video interventions' effects on respondents' written feedback, where 𝟙[Oppose] and 𝟙[Support] are indicator variables denoting that a respondent wrote opposing and supportive opinions toward the new regulation respectively. The video interventions consist of three main categories: nationalist videos (*Nationalist*), critical videos (*Critical*), and a mixed condition in which participants viewed both video types (*Mixed*). Both nationalist and critical videos were further subdivided into "emotional" (*Nat_Emotional* and *Crit_Emotional*) and "calm" versions (*Nat_Calm* and *Crit_Calm*). The Mixed group viewed emotional versions of both nationalist and critical videos. Controls include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to proposed changes to the Public Security Law. Robust t-statistics are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.10.

Table 3: Effect of Video Interventions on Respondents' Self- and Social Intolerance

	Self_Intolerance		**Social_Intolerance**	
	(1)	(2)	(3)	(4)
Nationalist	0.609***		0.612***	
	(3.553)		(3.283)	
Nat_Emotional		0.839***		1.178***
		(4.270)		(5.625)
Nat_Calm		0.359*		-0.002
		(1.761)		(-0.010)
Critical	-0.181		-0.353*	
	(-1.092)		(-1.861)	
Crit_Emotional		-0.258		-0.431*
		(-1.362)		(-1.946)
Crit_Calm		-0.094		-0.255
		(-0.488)		(-1.125)
Mixed	0.379**	0.384**	0.234	0.243
	(2.215)	(2.240)	(1.269)	(1.320)
Controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Obs	2,121	2,121	2,121	2,121
R-squared	0.053	0.056	0.064	0.076

Notes: This table shows the impact of video intervention on respondents' self-intolerance and expected social-intolerance. The video interventions consist of three main categories: nationalist videos (*Nationalist*), critical videos (*Critical*), and a mixed condition in which participants viewed both video types (*Mixed*). Both nationalist and critical videos were further subdivided into "emotional" (*Nat_Emotional* and *Crit_Emotional*) and "calm" versions (*Nat_Calm* and *Crit_Calm*). The Mixed group viewed emotional versions of both nationalist and critical videos. After the video intervention, respondents were asked whether they themselves would perceive someone opposing this new regulation as "unpatriotic" (*Self_Intolerance*), and were additionally required to predict whether others would hold such a perception (*Social_Intolerance*). All regression specifications include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to proposed changes to the Public Security Law. Robust t-statistics are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.10.

Table 4: The Impact of Self- and Social-Intolerance on Writing Comments

	𝟙[Oppose]				𝟙[Support]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self_Intolerance	-0.023***		-0.014***	-0.013***	0.005**		0.006**	0.006**
	(-8.069)		(-4.341)	(-4.096)	(2.148)		(2.496)	(2.287)
Social_Intolerance		-0.024***	-0.019***	-0.017***		-0.000	-0.003	-0.004**
		(-9.445)	(-6.671)	(-6.246)		(-0.180)	(-1.364)	(-2.010)
Nat_Emotional				-0.040*				0.081***
				(-1.857)				(3.489)
Nat_Calm				-0.018				-0.007
				(-0.764)				(-0.373)
Crit_Emotional				0.009				-0.006
				(0.341)				(-0.350)
Crit_Calm				0.109***				-0.002
				(3.696)				(-0.130)
Mixed				0.013				0.001
				(0.589)				(0.076)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2,121	2,121	2,121	2,121	2,121	2,121	2,121	2,121
R-squared	0.078	0.090	0.098	0.113	0.023	0.021	0.024	0.037

Notes: This table presents associations between self- and social-intolerance and respondents' decisions to submit written feedback. 𝟙[Oppose] and 𝟙[Support] indicate whether a respondent submitted opposing or supportive opinions, respectively. After the video interventions, respondents were asked whether they personally view opposition to the new regulation as "unpatriotic" (*Self_Intolerance*), and whether they believe others would hold such a view (*Social_Intolerance*). Columns 4 and 8 include treatment indicators: nationalist (*Nationalist*), critical (*Critical*), and mixed (*Mixed*) videos. Nationalist and critical conditions are further divided into "emotional" (*Nat_Emotional*, *Crit_Emotional*) and "calm" (*Nat_Calm*, *Crit_Calm*) versions. The mixed condition includes emotional versions of both. All regressions control for province fixed effects and demographics (age, gender, education, income, prior policy attention). Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Supplementary Experiment on Second Beliefs

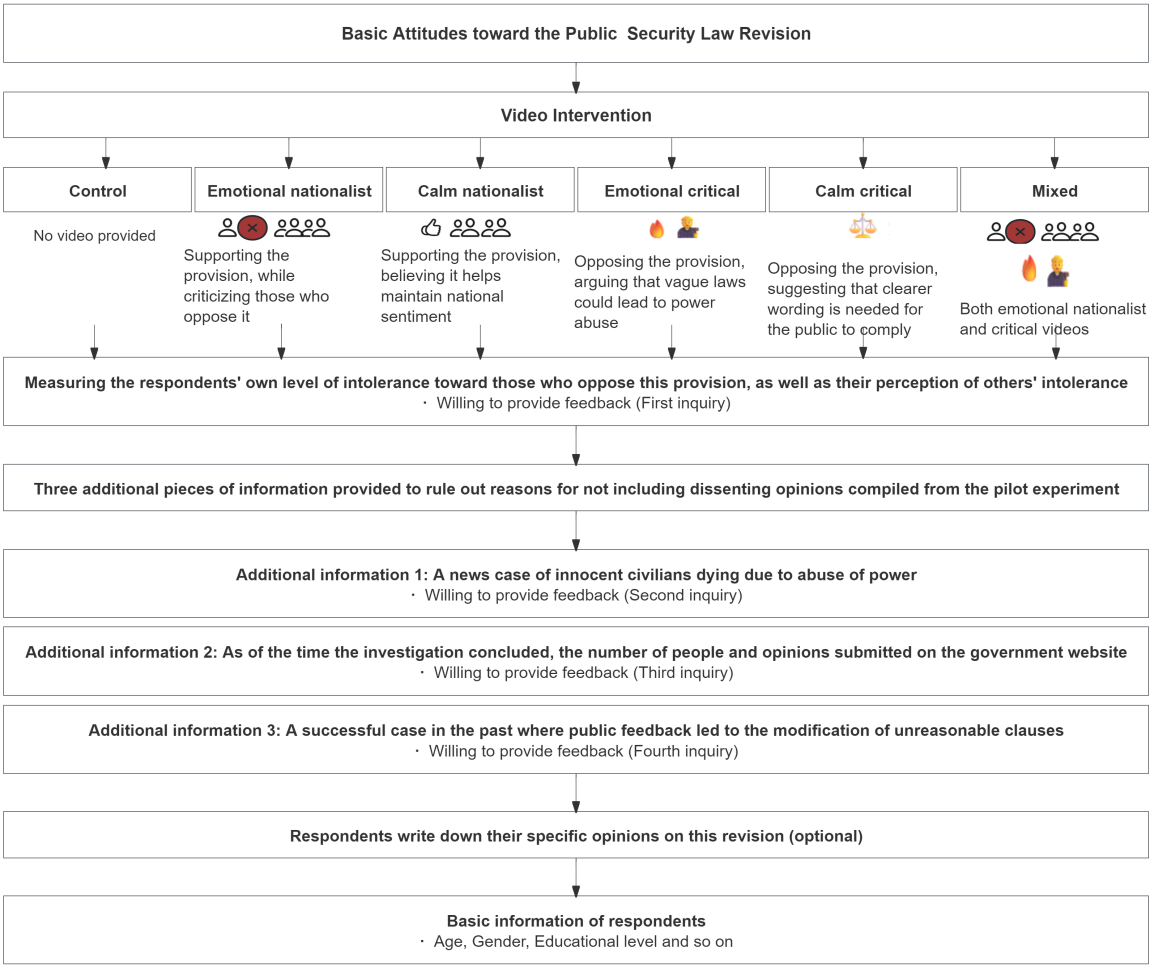
	𝟙[Oppose]				𝟙[Support]			
	Sup-Expt	Sup-Expt	Main-Expt	Pooled	Sup-Expt	Sup-Expt	Main-Expt	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nat_Emotional	0.004	-0.012	-0.046**	-0.045**	0.001	-0.004	0.087***	0.084***
	(0.123)	(-0.346)	(-2.043)	(-2.071)	(0.029)	(-0.169)	(3.543)	(3.488)
Informed				0.063**				0.020
				(2.156)				(0.907)
Nat_Emotional × Informed				0.047				-0.087***
				(1.174)				(-2.584)
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Province FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Obs	536	534	591	1,126	536	534	591	1,126
R-squared	0.000	0.061	0.086	0.058	0.000	0.056	0.062	0.035

Notes: This table presents the effects of informing respondents about the actual intolerance levels of others (denoted as $Informed$) on their propensity to express opposing (𝟙[Oppose]) or supportive (𝟙[Support]) opinions regarding the new regulation. In this supplementary experiment, respondents were exclusively exposed to nationalist video interventions. The variable $Nat_Emotional$ represents exposure to emotional nationalist videos. *Sup-Expt* refers to participants in the supplementary experiment in which participants viewed nationalist videos (emotional or calm) and were informed about others' true intolerance levels, while $Main\text{-}Expt$ denotes participants in the main experiment who viewed the nationalist videos. The *Pooled* sample combines nationalist video treatment groups from both experiments. All regression specifications include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to proposed changes to the Public Security Law. Robust t-statistics are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.10.

Appendix – For Online Publication Only

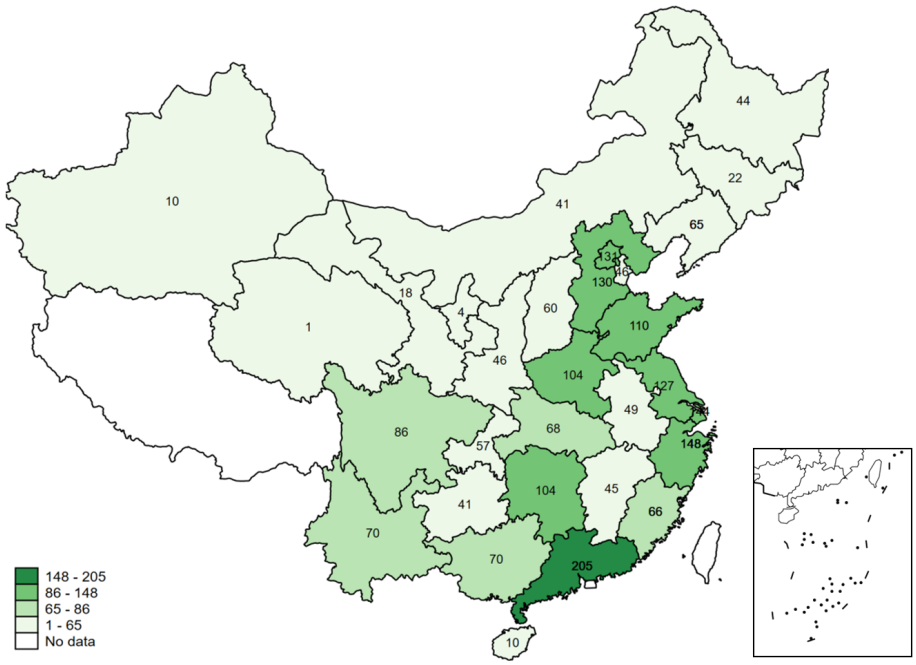
A Additional Information

Figure A1: Experiment Outline of the Main Experiment



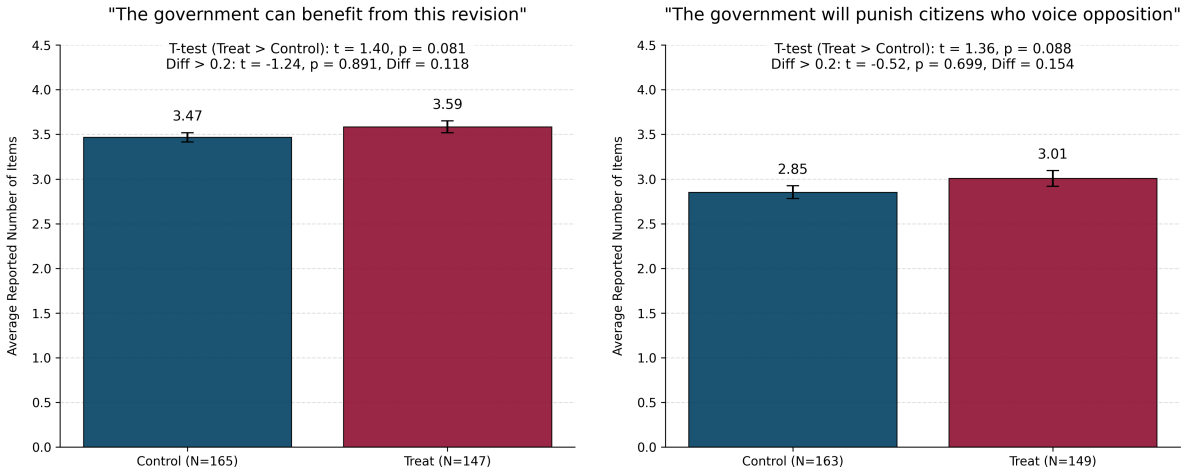
Notes: This figure illustrates the key experimental procedures implemented in the main experiment.

Figure A2: Distribution of Participants



Notes: This figure displays the geographical distribution of participants' home provinces.

Figure A3: List Experiment Results: Sensitivity of Political Expression



Notes: This figure shows results from two list experiments. The list experiments measure respondents' latent concerns about expressing political opinions. Results show small differences between the treatment and control groups. At the 5% significance level, we cannot reject the null hypothesis of no significant difference. Using 0.2 as the difference threshold, we can be 95% confident that the true difference does not exceed 0.2.

Table A1: The Content of Video Interventions and Information Shocks

Panel A: The Basic Content of Video Interventions		
Type	**Video Piece**	**Theme**
Nationalist	Emotional Nationalist	Uses forceful and emotional language and imagery to connect the legal revision with safeguarding national dignity.
	Calm Nationalist	Uses objective language and descriptions to explain the significance of this legal revision for national sentiment.
Critical	Emotional Critical	Uses intense and provocative language to critique the legal revision, linking it to the expansion and potential abuse of public power, emphasizing unclear and poorly vetted provisions.
	Calm Critical	Uses rational language to explain that legal provisions should be clarified to facilitate public understanding and compliance.
Panel B: The Basic Content of After-Video Informational Material Provided to All Respondents		
Information Shocks	**Theme**	**Main Content**
IS1	Legal revisions impact everyone	In 2014, a migrant worker, died during a police intervention in Taiyuan, reported as "migrant worker beaten to death over wage disputes." Excessive authority may harm innocent people; supervising laws is a citizen' s right and duty.
IS2	The actual number of commenter and comments	As of 4:00 PM on September 16, 2023, a total of 93,724 people have submitted 116,463 comments on the revision of this regulation via the official website, with numbers still increasing.
IS3	A precedent where public feedback successfully led to changes in a legal revision	The draft revision of the "Law of the People's Republic of China on the Protection of Women's and Children's Rights and Interests" was opened for public comment on December 24, 2021. Over 80,000 people submitted 420,000 comments, which successfully led to the removal of improperly or unreasonably worded provisions in the draft.

Table A2: Balance Test by Groups

	Control	Nationalist		Critical		Mixed
		Emotional	Calm	Emotional	Calm	
Sociodemographic Characteristics						
Gender	0.54 (0.03)	0.62** (0.03)	0.52 (0.03)	0.56 (0.03)	0.57 (0.03)	0.56 (0.02)
Age	3.39 (0.05)	3.55* (0.06)	3.40 (0.06)	3.48 (0.06)	3.39 (0.06)	3.31 (0.04)
Education	2.84 (0.04)	2.79 (0.04)	2.81 (0.04)	2.79 (0.04)	2.95 (0.03)	2.83 (0.03)
Income (yearly)	2.86 (0.05)	2.80 (0.06)	2.81 (0.07)	2.75 (0.06)	2.76 (0.06)	2.66 (0.04)
Attitudal Responses						
Patriotism	1.00 (0.03)	1.04 (0.03)	1.01 (0.04)	1.10 (0.04)	1.00 (0.04)	1.05 (0.03)
Cooperation	5.80 (0.14)	5.73 (0.15)	5.43 (0.17)	5.98 (0.16)	5.81 (0.15)	5.72 (0.11)
Money	7.47 (0.11)	7.47 (0.10)	7.58 (0.11)	7.55 (0.10)	7.71 (0.10)	7.53 (0.08)
Effort	7.66 (0.10)	7.59 (0.10)	7.67 (0.11)	7.43 (0.12)	7.34 (0.11)	7.80 (0.08)
Mistrust	6.32 (0.09)	6.48 (0.09)	6.42 (0.10)	6.52 (0.09)	6.31 (0.09)	6.38 (0.07)
Obs	359	313	279	294	283	594

Notes: This table presents the mean and standard error (in parentheses) of respondent characteristics for each subgroup. Kruskal-Wallis tests were conducted for continuous and ordinal variables due to non-normal distributions. The test for *attitude towards money* showed significant differences across groups (H=12.305, p=0.0308), but Dunn's post-hoc tests (Bonferroni-corrected) found no significant differences with the control group (all p-adj=1.0). The test for *Age* was marginally non-significant (H=10.122, p=0.0719). Chi-square tests were conducted for *Gender* (binary variable), with a significant difference between the *emotional nationalist* group and the control group (χ^2=4.730, p=0.0297). No other variables showed significant differences across groups (all p$\geq$0.1085). *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Impact of Video Intervention on Stated Willingness to Provide Feedback After Informational Interventions

	Willing_0			Willing_1			Willing_2			Willing_3		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nationalist	-0.127	-0.117		-0.149	-0.126		-0.189	-0.149		-0.155	-0.126	
	(-0.683)	(-0.630)		(-0.843)	(-0.705)		(-1.028)	(-0.803)		(-0.869)	(-0.700)	
Nat_Emotional			-0.423**			-0.200			-0.308			-0.158
			(-2.017)			(-0.984)			(-1.464)			(-0.759)
Nat_Calm			0.213			-0.047			0.022			-0.092
			(0.949)			(-0.217)			(0.098)			(-0.429)
Critical	1.094***	1.139***		0.810***	0.861***		0.926***	0.975***		0.802***	0.851***	
	(5.799)	(6.063)		(4.565)	(4.860)		(4.971)	(5.189)		(4.522)	(4.822)	
Crit_Emotional			0.772***			0.531**			0.686***			0.617***
			(3.510)			(2.529)			(3.112)			(2.982)
Crit_Calm			1.516***			1.205***			1.274***			1.095***
			(6.856)			(5.934)			(5.810)			(5.351)
Mixed	0.457**	0.484**	0.483**	0.371**	0.414**	0.416**	0.433**	0.446**	0.446**	0.331*	0.363**	0.365**
	(2.418)	(2.544)	(2.539)	(2.046)	(2.270)	(2.281)	(2.310)	(2.347)	(2.348)	(1.827)	(2.002)	(2.011)
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Province FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Obs	2,122	2,121	2,121	2,122	2,121	2,121	2,122	2,121	2,121	2,122	2,121	2,121
R-squared	0.029	0.055	0.063	0.020	0.046	0.050	0.024	0.058	0.062	0.019	0.056	0.058

Notes: This table presents the relationship between the video interventions and respondents' stated willingness to provide (oppositional) feedback across the three rounds of informational intervention. The video interventions consist of three main categories: nationalist videos (*Nationalist*), critical videos (*Critical*), and a mixed condition in which participants viewed both video types (*Mixed*). Both nationalist and critical videos were further subdivided into "emotional" (*Nat_Emotional* and *Crit_Emotional*) and "calm" versions (*Nat_Calm* and *Crit_Calm*). The Mixed group viewed emotional versions of both nationalist and critical videos. $Willing_n$ represents the respondent's willingness to provide feedback after receiving n rounds of informational interventions, where n=0 corresponds to the measurement taken immediately after watching the video. All regression specifications include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to the proposed changes to the Public Security Law. Robust t-statistics are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.10.

Table A4: Content of the Two List Experiments

List experiment 1: How many items do you support?		**List experiment 2: How many items concern you?**	
Control	**Treat**	**Control**	**Treat**
The government should increase investment in education to reduce educational inequality.	The government should increase investment in education to reduce educational inequality.	The leakage of personal privacy on social media is not effectively regulated by the government.	The leakage of personal privacy on social media is not effectively regulated by the government.
Environmental protection is everyone' s responsibility, and enterprises should take the lead.	Environmental protection is everyone' s responsibility, and enterprises should take the lead.	Risks of online scams and phishing websites are prevalent everywhere.	Risks of online scams and phishing websites are prevalent everywhere.
Smoking should be prohibited in all public places.	Smoking should be prohibited in all public places.	The government has installed excessive surveillance cameras in public spaces.	The government has installed excessive surveillance cameras in public spaces.
The government should raise the minimum wage based on regional characteristics.	The government should raise the minimum wage based on regional characteristics.	Credit card information is stolen during online shopping.	Credit card information is stolen during online shopping.
	The government can benefit from this revision		***The government would punish citizens who voice opposing opinions.***

Notes: This table presents the main content of the list experiments. The descriptions in bold are those seen only by the respondents in the treatment group. In each list experiment, respondents were asked to report the number of descriptions they agreed with, without specifying which ones they agreed with. This approach allows for the measurement of respondents' underlying attitudes.

Table A5: Effects of Stated Willingness to Provide Oppositional Feedback on Commenting Behavior

	𝟙[Oppose]					𝟙[Support]				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Willing_0$	0.032***				0.011***	0.001				-0.005*
	(11.170)				(3.231)	(0.280)				(-1.696)
$Willing_1$		0.033***			-0.002		0.003			-0.001
		(11.423)			(-0.678)		(1.566)			(-0.247)
$Willing_2$			0.037***		0.005			0.005**		0.001
			(12.884)		(1.096)			(2.228)		(0.364)
$Willing_3$				0.044***	0.034***				0.007***	0.010***
				(15.357)	(7.704)				(3.253)	(2.680)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2,121	2,121	2,121	2,121	2,121	2,121	2,121	2,121	2,121	2,121
R-squared	0.123	0.118	0.145	0.172	0.179	0.021	0.022	0.023	0.026	0.029

Notes: This table presents the correlation between respondents' stated willingness to provide (oppositional) feedback and choosing to provide actual feedback. We measured respondents' willingness to express opinions after each video intervention and following every textual information intervention. $Willing_n$ represents the respondent's willingness to provide feedback after receiving n rounds of informational interventions, where n=0 corresponds to the measurement taken immediately after watching the video. All regression specifications include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to the proposed changes to the Public Security Law. 𝟙[Oppose] and 𝟙[Support] are indicator variables denoting that a respondent provided opposing and supportive opinions, respectively, toward the Public Security Law. All regression specifications include province fixed effects and control for demographic characteristics including age, gender, education level, income, and prior attention to proposed changes to the Public Security Law. Robust t-statistics are reported in parentheses, *** p<0.01, ** p<0.05, * p<0.10.

Table A6: Respondents' Self- and Social-Intolerance after 6 and 12 months

Group	Self Intolerance			Social Intolerance		
	Baseline	Survey after 6 months	Survey after 12 months	Baseline	Survey after 6 months	Survey after 12 months
Nationalist	5.53	4.06	3.83	6.11	4.58	4.05
Critical	4.71	4.07	3.76	5.21	4.39	4.09
Mixed	5.27	4.12	3.75	5.77	4.76	4.11
Control	4.95	4.09	3.79	5.62	4.72	4.14

Notes: This table presents the changes in self- and social-intolerance among respondents from the main experiment who participated in the 6-month and 12-month follow-up surveys. $Self_Intolerance$ measures whether respondents themselves perceive opponents of the new regulation as unpatriotic, while $Social_Intolerance$ captures their beliefs about whether others hold such perceptions.

Table A7: Recall and Attitudes for 6 and 12 Month Follow-up Groups

Panel A: Own recall and beliefs about the law

Group	Remember exactly what the new law is about	Remember the video content	Willing to give comments again	Opinion on the law is similar to initial opinion
Nationalist	11.49%	19.54%	16.09%	21.43%
Critical	3.33%	16.67%	16.67%	33.33%
Mixed	13.64%	32.95%	15.91%	7.14%
Control	10.29%	-	17.65%	25.00%

Panel B: Changes in respondents' beliefs of others' attitudes

Group	Belief about the average proportion of the public that supports the new law (%)		Average perception of difference between support and opposition (%)	
	Main Expt	Follow-up	Main Expt	Follow-up
Nationalist	74.08	66.6	12.94	18.05
Critical	61.67	63.4	20.46	19.01
Mixed	65.01	67.58	13.01	14.92
Control	69.15	70.28	13.40	14.24

Notes: The table presents the beliefs and attitudes of respondents who participated in the main experiment, during follow-up surveys at 6 and 12 months. Panel A shows information on respondents' memory of the video content, the legal provisions, and whether their willingness to express an opinion changed. Panel B displays respondents' perceptions of others' attitudes toward the law, compared to the opinions they expressed in the main experiment.

Table A8: The Content of Common Videos Related to the Public Security Law

Total Video Views (Thousands), 2023.9.1–2023.9.30				
	TikTok (Douyin)		Kuaishou	
Video Type	**Emotional**	**Calm**	**Emotional**	**Calm**
Nationalist	18,885	4,709	900	105
Critical	1,459	936	342	28

Notes: This table shows the total views of a subset of short videos related to the Public Security Law that appeared on during the public consultation period on China's top two short video platforms, TikTok (Douyin) and Kuaishou. Since the total view count is unavailable, we use the total number of "likes" and apply a conservative 5% like-to-view ratio to estimate the total views.